Clinical exome sequencing for cerebellar ataxia and spastic paraplegia uncovers novel gene-disease associations and unanticipated rare disorders

De Warrenburg, Bart P. van; Schouten, Meyke I.; de Bot, Susanne T.; Vermeer, Sascha; Meijer, Rowdy; Pennings, Maartje; Gilissen, Christian; Willemsen, Michel A. A. P.; Scheffer, Hans; Kamsteeg, Erik-Jan

Published in:
European Journal of Human Genetics

DOI:
10.1038/ejhg.2016.42

Publication date:
2016

European Journal of Human Genetics (2016) 24, 1460–1466

www.nature.com/ejhg

EJHG Open

ARTICLE

Clinical exome sequencing for cerebellar ataxia and spastic paraplegia uncovers novel gene–disease associations and unanticipated rare disorders

Bart P van de Warrenburg[1], Meyke I Schouten[2], Susanne T de Bot[3], Sascha Vermeer[3], Rowdy Meijer[2], Maartje Pennings[2], Christian Gilissen[2], Michèl AAP Willemsen[1], Hans Scheffer[2] and Erik-Jan Kamsteeg*[,2]

Cerebellar ataxia (CA) and hereditary spastic paraplegia (HSP) are two of the most prevalent motor disorders with extensive locus and allelic heterogeneity. We implemented clinical exome sequencing, followed by filtering data for a 'movement disorders' gene panel, as a generic test to increase variant detection in 76 patients with these disorders. Segregation analysis or phenotypic re-evaluation was utilized to substantiate findings. Disease-causing variants were identified in 9 of 28 CA patients, and 8 of 48 HSP patients. In addition, possibly disease-causing variants were identified in 1 and 8 of the remaining CA and HSP patients, respectively. In 10 patients with CA, the total disease-causing or possibly disease-causing variants were detected in 8 different genes, whereas 16 HSP patients had such variants in 12 different genes. In the majority of cases, the identified variants were compatible with the patient phenotype. Interestingly, in some patients variants were identified in genes hitherto related to other movement disorders, such as *TH* variants in two siblings with HSP. In addition, rare disorders were uncovered, for example, a second case of HSP caused by a *VCP* variant. For some patients, exome sequencing results had implications for treatment, exemplified by the favorable L-DOPA treatment in a patient with HSP due to *ATP13A2* variants (Parkinson type 9). Thus, clinical exome sequencing in this cohort of CA and HSP patients suggests broadening of disease spectra, revealed novel gene–disease associations, and uncovered unanticipated rare disorders. In addition, clinical exome sequencing results have shown their value in guiding practical patient management.
European Journal of Human Genetics (2016) **24,** 1460–1466; doi:10.1038/ejhg.2016.42; published online 11 May 2016

INTRODUCTION

Massive parallel sequencing is rapidly being applied to identify disease-causing variants in disorders with locus and allelic heterogeneity. Both targeted gene panel sequencing[1] and exome or genome sequencing approaches[2–4] are being used in diagnostic neurogenetics. Both approaches allow for parallel analyses of multiple genes (panels) associated with a given disorder in a single test. However, where targeted gene panel approaches are limited to this, exome (or genome) sequencing enables variant detection beyond a predefined gene panel or in 'broader' gene panels. The unbiased nature of exome sequencing can thus result in the discovery of novel gene–disease associations, in the expansion of disease spectra, or in the recognition of very rare and unanticipated disorders. On the other hand, exome sequencing bears the risk of finding pathogenic variants in genes not related to the original request. To minimize the chance of having such 'unsolicited' findings, bioinformatic gene panels may be used for initial filtering.[5]

Here, we have used clinical exome sequencing with bioinformatic gene panel filtering for variant detection in cerebellar ataxia (CA) and hereditary spastic paraplegia (HSP). The prevalence rates of CA and HSP are 1–5 per 100 000[6] and 1–10 per 100 000,[7] respectively. CA and HSP are genetically heterogeneous, with over 70 loci/genes involved in each, often with little or no clues for genetic differentiation. In this study, the exome sequencing data were managed using a bioinformatic gene panel filter containing ~200 genes for central motor disorders, including not only CA and HSP, but also Parkinson's disease and various hyperkinetic movement disorders.

SUBJECTS AND METHODS

Patients

In total, 76 patients with CA ($n=28$) or HSP ($n=48$) were included for clinical exome sequencing. The cohort was largely nonconsecutive. Patients with a visiting history from our center were selected from corresponding databases, contacted, and then referred to and counseled by a clinical geneticist. New patients were recruited at their first visit. No restrictions on age at onset or inheritance pattern were made, but there had to be a clinical suspicion of a genetic etiology, based on the combination of a CA and/or HSP phenotype plus either an onset age below 45 years or a positive family history. The latter was liberally defined as any affected family member suggestive of dominant, recessive, or X-linked inheritance. Moreover, mutations in the most likely gene(s) had to be excluded. All patients were counseled by a clinical geneticist and informed consent was obtained. From each family, only one patient (proband) was included for exome sequencing. Candidate mutations were verified by Sanger sequencing in affected family members whenever possible. All but three patients were adults at the time of testing, but in many the disease had started in childhood (see Tables 1 and 2). Most patients had already undergone extensive genetic testing, with a mean of 4.5 genes for each group. Autosomal dominant cerebellar ataxias caused by repeat expansions were excluded in all CA patients with dominant or sporadic modes of inheritance. In addition, SPG types 4 and 7

[1]Department of Neurology, Radboud University Medical Center, Nijmegen, The Netherlands; [2]Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands; [3]Department of Human Genetics, University Medical Center Groningen, Groningen, The Netherlands
*Correspondence: Dr E-J Kamsteeg, Department of Human Genetics, Radboud University Medical Center, Geert Grooteplein Zuid 10, 6500 HB Nijmegen, The Netherlands. Tel: 31 24 3617757; Fax: 31 24 3616658; E-mail: erik-jan.kamsteeg@radboudumc.nl
Received 14 September 2015; revised 25 March 2016; accepted 12 April 2016; published online 11 May 2016

were excluded in ~75 and 50% of the HSP patients before exome sequencing. The set of genes excluded per patient was quite variable, but this reflects the variation of clinical practice in case of different requesting and referring physicians. Clinical exome sequencing was approved by the Medical Review Ethics Committee, Region Arnhem-Nijmegen, Number 2011/188.

Exome sequencing and data analysis

Exome sequencing was essentially performed as previously described.[5] Briefly, capture of exons was done using an Agilent SureSelect Human All Exon 50 Mb Kit (Santa Clara, CA, USA). Sequencing was performed using a Life Technologies 5500XL machine (Thermo Fisher, Waltham, MA, USA) or an Illumina Hiseq 2000 (San Diego, CA, USA). Read mapping and variant calling were done using LifeScope Life Technologies (Thermo Fisher) for the 5500XL data or BWA (mapping) and GATK (calling) for the Illumina data. A filter for a 'movement disorders' gene panel was applied. This panel consists of ~200 genes implicated in various forms of CA, HSP, Parkinson's disease, genetic choreas, and other hyperkinetic movement disorders. Only genes with substantial evidence (multiple families, functional evidence, and/or literature reports) were included in this panel. Genes with repeat expansions as molecular mechanism only were not included either. The genes in this panel and coverage statistics are provided as Supplementary File 1) and can also be found at www.genomediagnosticsnijmegen.nl/exome. Variants were prioritized based on the following criteria: frequency (<5% dbSNP, <1% in-house database of >5000 exomes), nucleotide and amino acid conservation (based on alignments), relation of the gene to disease (per family), and inheritance pattern.[5] All reported variants were confirmed by Sanger sequencing.

All possible causative variants have been deposited into the appropriate gene-specific databases (https://grenada.lumc.nl/LOVD2/mendelian_genes (Patient ID 0064817), PMM2 (Patient ID 0064818) or http://databases.lovd.nl/shared/variants (SPG7), 00054866 (GOSR2), 00054867 (SCN8A), 00054869 (C10orf2), 00054870 (KCNC3), 00054871& 00054875 (OPA1), 00054872 (GBA2), 00054873 (FA2H), 00054874 (KIAA0196), 00054876 (TH), 00054877 and 00054878 (KIF1A), and 00054879 (LRRK2).

RESULTS

Diagnostic yield of exome sequencing for CA and HSP

Initially, variants were identified by exome sequencing and an in-house designed 'movement disorder' filter in 33 of 76 patients. Some of those were directly considered causative, whereas others needed follow-up by segregation analysis, clinical re-evaluation, or biochemical tests. Pathogenicity was based on nature of the variants and previous reports (see Tables 1 and 2) and frequency in controls and prediction algorithms (Table 3). Of the 28 CA patients, eventually 9 had causative variants, 1 a possibly causative variant, and another 4 likely noncausative variants (Table 4). Of the 48 HSP patients, 8 had causative, 8 possibly causative, and 3 had likely noncausative variants. Thus, the diagnostic yield of clinical exome sequencing in this cohort is somewhere between 17 and 26 of 76 (~1 in 4).

Molecular diagnoses in cerebellar ataxia

Three patients with autosomal recessive or sporadic CA had biallelic *SPG7* variants (Table 1a). Spastic paraplegia type 7 (SPG7) accounts for 5–12% of autosomal recessive forms of HSP, depending on the population tested.[20] In approximately half of the SPG7 patients, CA is an additional clinical feature.[21] However, two of the three patients described here had no pyramidal signs, but rather a pure CA; the third did have pyramidal features at clinical re-evaluation. Two of these patients are compound heterozygous for two common variants in spastic paraplegia type 7, p.(Arg485_Glu487del) and p.(Ala510Val), suggesting that pure ataxia caused by *SPG7* variants is not caused by specific variants, but part of the *SPG7* disease spectrum.

Three other patients (4–6; Table 1a) had ataxia complicated by seizures, and exome sequencing revealed variants in genes involved in rare disorders. Patient 4 carried a *de novo* variant in *SCN8A*. Variants in this gene are associated with early epileptic infantile encephalopathy type 13.[22] Patient 5 was homozygous for the only *GOSR2* variant known to date to cause progressive myoclonic epilepsy type 6,[23] whereas patient 6 has an earlier described nonsense variant in *CACNA1A*, a gene involved in various disorders with ataxia and/or epilepsy.

The remainder of the variants (Table 1a) was considered to be pathogenic in CA for several reasons. Patient 7 has ataxia complicated

Table 1 Cerebellar ataxia patients with (a) causative variants and (b) possible causative variants identified by exome sequencing

Patient	*Sex*	*Onset*	*Additional features*	*Inherit.*	*Gene*	*Variant (DNA) (protein)*	*Zygosity*	*Disorder (inherit.)*
(*a*)[a]								
1	F	58 y	None	AR	*SPG7*	c.1454_1462del; p.(Arg485_Glu487del)[17] c.1529C>T; p.(Ala510Val)[18]	ch	SPG7 (AR)
2	M	38 y	None	AR	*SPG7*	c.1454_1462del; p.(Arg485_Glu487del)[17] c.1529C>T; p.(Ala510Val)[18]	ch	SPG7 (AR)
3[b]	M	52 y	Mild pyramidal signs LL	S	*SPG7*	c.1529C>T; p.(Ala510Val)[18] c.1756G>T; p.(Glu586*)	ch	SPG7 (AR)
4	M	6 mo	ID, spasticity, epilepsy	S	*SCN8A*	c.4351G>A; p.(Gly1451Ser)[16]	*dn*	EIEE13 (AD)
5	F	2 y	Myoclonus, dystonia, epilepsy, high CK	AR	*GOSR2*	c.430G>T; p.(Gly144Trp)[c,9]	ho	EPM6 (AR)
6	F	12 y	Epilepsy	AD	*CACNA1A*	c.5569C>T; p.(Arg1857*)[8]	he	EA2 (AD)
7	F	<3 y	Ovarian dysgenesis, deafness, neuropathy, ID	AR	*C10orf2*	c.1306G>C; p.(Gly436Arg)[c]	ho	MTDS7 (AR)
8	F	<10 y	Neuropathy	AR	*PMM2*	c.422G>A; p.(Arg141His)[14] c.722G>C; p.(Cys241Ser)[15]	ch	CDG1a (AR)
9[b]	M	1 y	Pyramidal signs LL	S	*KCNC3*	c.1268G>A; p.(Arg423His)[10]	he	SCA13 (AD)
(*b*)[d]								
10a	F	U	Cataract, neuropathy	U	*OPA1*	c.965 T>G; p.(Met322Arg)	he	DOA+(AD, AR)

Abbreviations: F, female; M, male; LL, lower limbs; ID, intellectual disability; R, autosomal recessive; D, autosomal dominant; S, sporadic; U, unknown because of adoption; y, year; mo, month; ch, compound heterozygous; *dn*, *de novo*; he, heterozygous; ho, homozygous.
Gene names are according to the HUGO Gene Nomenclature Committee (HGNC) and variant nomenclature is according to the Human Genome Variation Society (HGVS) guidelines.
[a]Reference sequences used are: NM_003119.2 (*SPG7*); NM_014191.2 (*SCN8A*); NM_054022.2(*GOSR2*); NM_001127221.1 (*CACNA1A*); NM_021830.4 (*C10Orf2*); NM_000303.2 (*PMM2*); and NM_004977.2 (*KCNC3*). OMIM disease IDs used are: 607259 (SPG7); 614558 (EIEE13); 614018 (EPM6); 108500 (EA2); 271245 (MTDS7); 212065 (CDG1a); and 605259 (SCA13).
[b]Affected family member(s) with same genotype.
[c]Same patient described before.[5]
[d]Reference sequence used are: NM_130837.2 (*OPA1*). OMIM disease ID used: 125250 (DOA+).

Table 2 Hereditary spastic paraplegia patients with (a) causative variants and (b) possible causative variants identified by exome sequencing

Patient	*Sex*	*Onset*	*Additional features*	*Inherit*	*Gene*	*Variant (DNA) (protein)*	*Zygosity*	*Disorder (inherit.)*
(a)[a]								
11	M	<20 y	None	AD	*RTN2*	c.939del; p.(Thr314fs)[b]	he	SPG12 (AD)
12[c]	M	<3 mo	None	AD	*KIF5A*	c.611G>A; p.(Arg204Gln)[b,11]	he	SPG10 (AD)
13[c]	F	47 y	None	AD	*KIF5A*	c.396G>A; r.396_397ins247[b]	he	SPG10 (AD)
14	M	29 y	Neuropathy	AD	*KIF5A*	c.751G>A; p.(Glu251Lys)[b,11]	he	SPG10 (AD)
15	M	<3 y	Ataxia, cataract, neuropathy, mild ID, thin corpus callosum, white matter changes	AR	*GBA2*	c.2233C>T; p.(Gln745*)	ho	SPG46 (AR)
16[c]	M	4 y	Ataxia, hypomimia	AR	*FA2H*	c.554G>A; p.(Trp185*) c.21del; p.(Ala8fs)[d]	ch	SPG35 (AR)
17	M	<1 y	Parkinsonism, ID, slow vertical saccades	U	*ATP13A2*	c.2675G>A; p.(Gly892Asp)[b]	ho	PARK9 (AR)
18[c]	F	15 y	None	AD	*KIAA0196*	c.3103C>T; p.(Arg1035Cys)[b]	he	SPG8 (AD)
(b)[e]								
19	M	6 y	None	S	*VCP*	c.278G>A; p.(Arg93His)[b,f]	he	IBMPFD (AD)
20[c]	M	20 y	None	AR	*TH*	c.772G>A; p.(Glu258Lys)	ho	THD (AR)
21	F	U	None	AD	*LRRK2*	c.2500+5_ 2500+8del r.2311_2500del/r.2345_2500del[12]	he	PARK8 (AD)
22[c]	M	<3 mo	None	AD	*KIF1A*	c.221A>G; p.(Tyr74Cys)	he	SPG30 (AR)
23[c]	M	<20 y	None	AD	*KIF1A*	c.1894C>T; p.(Gln632*)	he	SPG30 (AR)
24	M	53 y	None	AD	*KIAA0196*	c.247G>A; p.(Glu83Lys)	he	SPG8 (AD)
25	M	<5 y	Motor neuropathy, ID, mild ataxia	S	*AFG3L2*	c.2314C>T; p.(Leu772Phe)	ho	SPAX5 (AR)
26	M	20 y	None	AD	*OPA1*	c.113_130del; p.(Arg38_Ser43del)[13]	he	DOA+(AD, AR)

[a]Reference sequences used are: NM_005619.3 (*RTN2*); NM_004984.2 (*KIF5A*); NM_020944.2 (*GBA2*); NM_024306.4 (*FA2H*); NM_022089.2 (*ATP13A2*); and NM_014846.3 (*KIAA0196*). OMIM disease IDs used are: 604805 (SPG12); 604187 (SPG10); 614409 (SPG46); 612319 (SPG35); 606693 (PARK9); and 603563 (SPG8).[5]
[b]Same patient described before.[5]
[c]Affected family member(s) with same genotype.
[d]Detected after Sanger sequencing only.
[e]Reference sequences used are: NM_007126.3 (*VCP*); NM_199292.2 (*TH*) NM_198578.3 (*LRRK2*); NM_001244008.1 (*KIF1A*); NM_014846.3 (*KIAA0196*); NM_006796.2 (*AFG3L2*); and NM_130836.2 (*OPA1*). OMIM disease IDs used are: 167320 (IBMPFD); 605407 (tyrosine hydroxylase deficiency (THD)); 607060 (PARK8); 610357 (SPG30); 603563 (SPG8); 604581 (SPAX5); and 125250 (DOA+).
[f]Variant from unaffected father, but IBMPFD shows reduced penetrance for most features.

by ovarian dysgenesis, deafness, neuropathy, and a mild intellectual disability that is compatible with mitochondrial DNA depletion syndrome caused by *C10orf2* (*Twinkle*) variants. In patient 8, two well-characterized *PMM2* variants with reduced enzymatic activity[24] were identified. Despite the often severe CDG1a phenotype associated with *PMM2* variants, a relative mild neurological phenotype with ataxia as the main feature has been described before.[25,26] The missense variant in *KCNC3* (patient 9) was considered causative, as it has been described earlier in ataxia and, in the present case, was found in an affected sibling and in a mosaic state in the unaffected mother.

A heterozygous variant in *OPA1*, the gene for dominant optic atrophy type 1 and optic atrophy-plus syndrome, was considered to be possibly causative (patient 10, Table 1b). However, despite segregation of this variant with the disease in two siblings and despite the fact that this variant affects an evolutionary conserved residue, it is quite uncertain whether this variant is the cause of the patient's ataxia. The optic atrophy plus syndrome is mostly a recessive disorder, and no second variant was identified using this test. The test has several limitations, including incomplete coverage of the gene, no detection of deep-intronic variants (as have been found in OPA1[27]), or the poor detection of exon deletions/duplications. In addition, the *OPA1* gene seems to be quite a polymorphic gene, with many nonpathogenic sequence variants. Analysis of the *OPA1* gene from our anonymized cohort of healthy parents from trio-based studies ($n=2224$) revealed 7 infrequent (<1% carrier frequency) sequence variants that lead to substitutions of moderately or well conserved amino acids. Thus, further testing of the *OPA1* gene and its involvement in this family's phenotype is warranted. Unfortunately, further segregation analysis was not possible as the parents were not available for testing.

Other variants were thought to be possibly causative initially, but eventually considered likely noncausative, for the following reasons: (1) they did not segregate with the disease or (2) no second pathogenic variant in the same gene was detected by Sanger sequencing in recessive disorders (Supplementary File 2). Of note, *AP5Z1 (KIAA0415)* variants have been described in autosomal recessive spastic paraplegia type 48.[28] However, only two families have been described in the original paper, and in a replication study no further pathogenic *AP5Z1* variants were identified in a cohort of 127 ataxia patients.[29] Two affected siblings from the original study were homozygous for a frameshift variant, whereas a third unrelated patient had only one heterozygous frameshift variant. These findings are inconclusive concerning the putative inheritance pattern proposed for *AP5Z1* variants. The *AP5Z1* nonsense variant in patient 28 was not present in his affected sibling, making it an unlikely cause of the CA in both siblings. Furthermore, using exome sequencing, we recently identified another *AP5Z1* heterozygous null variant (c.1312-2A>G) that did not segregate with spastic ataxia in another family (not part of this cohort), whereas a homozygous *AP5Z1* nonsense variant was revealed in a patient with retinitis pigmentosa but no motor symptoms. Therefore, it seems unlikely that *AP5Z1* variants are related to disease and we consider SPG48 for now to be unconfirmed.

Molecular diagnoses in hereditary spastic paraplegia

In eight HSP patients, clinical exome sequencing revealed variants that were considered causative because they had been described in other

Table 3 Frequency and pathogenicity scores of identified causative and possible causative variants

Gene	*Variant (DNA)*	*Variant (protein) and reference*	*Allele frequency in controls*	*Number of heterozygous controls*	*Number of homozygous controls*	*SIFT*	*Polyphen (HumVar)*	*Align GVGD*
AFG3L2	c.2314C>T	p.(Leu772Phe)	0.0026 (Eu)	19	0	0.23	**1.000**	C0
ATP13A2	c.2675G>A	p.(Gly892Asp)	0.00003 (Eu)	2	0	**0.02**	**0.947**	C0
C10orf2	c.1306G>C	p.(Gly436Arg)[5]	Absent	0	0	**0**	**1.000**	**C65**
CACNA1A	c.5569C>T	p.(Arg1857*)[8]	Absent	0	0	—	—	—
FA2H	c.554G>A	p.(Trp185*)	Absent	0	0	—	—	—
	c.21del	p.(Ala8fs)	Absent	0	0	—	—	—
GBA2	c.2233C>T	p.(Gln745*)	Absent	0	0	—	—	—
GOSR2	c.430G>T	p.(Gly144Trp)[9]	0.00007 (Eu)	5	0	**0**	**1.000**	**C65**
KCNC3	c.1268G>A	p.(Arg423His)[10]	Absent	0	0	**0**	**1.000**	C25
KIAA0196	c.247G>A	p.(Glu83Lys)	Absent	0	0	**0.03**	**0.982**	C0
	c.3103C>T	p.(Arg1035Cys)	Absent	0	0	**0.01**	0.751	C0
KIF1A	c.221A>G	p.(Tyr74Cys)	Absent	0	0	**0**	**0.999**	**C65**
	c.1894C>T	p.(Gln632*)	Absent	0	0	—	—	—
KIF5A	c.611G>A	p.(Arg204Gln) [11]	Absent	0	0	**0.03**	**1.000**	C0
	c.396G>A	r.396_397ins247[5]	Absent	0	0	—	—	—
	c.751G>A	p.(Glu251Lys) [11]	Absent	0	0	**0**	**1.000**	**C55**
LRRK2	c.2500+5_2500+8del	r.2311_2500del/ r.2345_2500del[12]	0.0001 (EA)	1	0	—	—	—
OPA1	c.113_130del	p.(Arg38_Ser43del)[13]	0.0022 (All)	20	1	—	—	—
	c.965 T>G	p.Met322Arg	Absent	0	0	**0.02**	0.809	C0
PMM2	c.422G>A	p.(Arg141His)[14]	0.011 (All)	50	0	**0**	0.510	C25
	c.722G>C	p.(Cys241Ser)[15]	0.00058 (Af)	9	0	**0.04**	**0.982**	C0
RTN2	c.939del	p.(Thr314fs) [5]	Absent	0	0	—	—	—
SCN8A	c.4351G>A	p.(Gly1451Ser)[16]	Absent	0	0	**0**	**0.954**	C0
SPG7	c.1454_1462del	p.(Arg485_Glu487del)[17]	0.0015 (All)	6	6	—	—	—
	c.1529C>T	p.(Ala510Val)[18]	0.004 (All)	45	0	**0**	**1.000**	**C65**
	c.1756G>T	p.(Glu586*)	Absent	0	0	—	—	—
TH	c.772G>A	p.(Glu258Lys)[5]	0.005 (SA)	46	1	**0**	0.669	**C55**
VCP	c.278G>A	p.(Arg93His)[5]	0.000015 (Eu)	1	0	0.12	**0.991**	C0

Abbreviations: EA, European American (from EVS (Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) May 2015); AA, African American (from EVS); All, EA and AA (from EVS); Eu, European (from ExAC (Exome Aggregation Consortium (ExAC), Cambridge, MA (URL: http://exac.broadinstitute.org) 11 May 2015); SA, South Asian (from ExAC); Af, African (from ExAC).
Gene names are according to the HUGO Gene Nomenclature Committee (HGNC) and variant nomenclature is according to the Human Genome Variation Society (HGVS) guidelines.
Sorting intolerant from tolerant (SIFT) scores range from 0 to 1. The amino acid substitution is predicted damaging if the score is $\leq$0.05, and tolerated if the score is $>$0.05.[19] PolyPhen-2 (Polymorphism Phenotyping v2) scores range from 0 to 1. The amino acid substitution is predicted benign if the score is $\leq$0.2, possibly damaging if the score is 0.2–0.85, or probably damaging if the score is $>$0.85. Align GVGD provides a series of ordered grades ranging from the most likely deleterious 'C65' to the least likely deleterious 'C0', and has 7 grades: C65, C55, C45, C35, C25, C15, and C0. Deleterious scores are in bold.

Table 4 Diagnostic yield of gene panel analysis after exome sequencing

Disease group	*Causative variants*	*Possibly causative variants*	*Less likely causative variants*	*No causative variants*	*Total*
CA	9	1	4	14	28
HSP	8	8	3	29	48
Total	17	9	7	43	76

Numbers of CA or HSP patients are categorized according to findings in clinical exome sequencing after gene panel analyses.

HSP patients before, they were null variants, or they segregated with the disease (patients 11–18, Table 2a). Three of those were in *KIF5A*, indicating that SPG10 is relatively frequent in this cohort of HSP patients.

Surprisingly, one of the *KIF5A* variants (detected in patient 13 and her affected father) leads to intron retention in the mature mRNA in lymphocytes (Supplementary File 3). This is predicted to result in a premature termination codon. Despite the suggestion of gain-of-function or dominant-negative effects of variants in SPG10,[30] the variant spectrum consists of missense, splice site, and frameshift variants, indicating that *KIF5A* haploinsufficiency may also be a mechanism underlying SPG10.

Patient 11 has a heterozygous frameshift variant in *RTN2*. As two of the three SPG12-causing variants in this gene described to date are null variants,[31] haploinsufficiency of *RTN2* seems to cause SPG12. Patient 15 has a homozygous nonsense variant in *GBA2*. This patient has HSP complicated with a mild intellectual disability, cataract, and neuropathy that are major characteristics of SPG46 caused by *GBA2* variants.

Genetic and clinical follow-up of two other patients also confirmed the role of the identified variants in their spastic syndrome. In patient 16, one heterozygous *FA2H* variant was complemented with a second variant identified by Sanger sequencing. The latter was missed by exome sequencing because of coverage issues. The MRI in this patient did not show iron accumulation, a variable finding in SPG35 due to *FA2H* variants. 'Reverse phenotyping' was convincing for Kufor–Rakeb syndrome in patient 17 with *ATP13A2* variants, because of the additional presence of mild Parkinsonism, upgaze limitation, and

cognitive decline (see Supplementary File 4). A pure HSP in patient 18 and her affected father is likely to be due to a *KIAA0196* variant.

A relatively large group of HSP patients had variants of uncertain pathogenicity (Table 2b). Variants in *VCP* can cause amyotrophic lateral sclerosis type 14 (ALS14) and inclusion body myopathy with early-onset Paget disease and frontotemporal dementia type 1 (IBMPFD1), a very heterogeneous disorder with reduced penetrance of each feature. Recently, we reported on a *VCP* variant as a new cause of HSP.[32] The p.(Arg93His) variant in *VCP* in patient 19 is novel. However, another variant of the same codon, p.(Arg93Cys), has been described in IBMPFD1,[33] suggesting that patient 19 is the second case of HSP caused by a *VCP* variant. Similarly, patient 20 and his sibling have HSP and a homozygous variant in *TH* that is normally involved in autosomal recessive dopamine-responsive dystonia. The p.(Glu258Lys) variant in *TH* in this family is likely to be causative, as the same change, but of the adjacent and equally conserved codon (p.(Glu259Lys)), has been described in patients with dopamine-responsive dystonia.[34] Unfortunately, a lumbar punction to study dopamine metabolites and a trial with L-DOPA was refused by both siblings. Interestingly, a variant in the *GCH1* gene, causing autosomal dominant dopamine-responsive dystonia, has recently been associated with dopamine-responsive, autosomal dominant HSP,[35] suggesting a link between HSP and dystonia.

In the same group of possibly causative variants in HSP, several variants in genes implicated in HSP (*KIAA0196* and *KIF1A*) or in other central motor disorders (*LRRK2*, *AFG3L2*, *OPA1*) were identified, and some of the variants have been published before (but not in HSP). Segregation analyses or functional studies were not possible to confirm or exclude the role of these variants in HSP because of various reasons. However, an interesting observation is the finding of two heterozygous *KIF1A* variants segregating in two families with autosomal dominant HSP. *KIF1A* variants have been associated with autosomal recessive HSP in two families,[36] and also in *de novo* autosomal dominant intellectual disability with motor symptoms.[37,38] All variants described in *KIF1A* in these conditions affect the kinesin motor domain. The mechanism underlying these two different disorders has not been established, but it is tempting to speculate that the *KIF1A* variant that segregates with HSP in the family of patient 22, which is also in its kinesin motor domain, may cause autosomal dominant HSP. An apparent dominant inheritance of HSP with normal cognition was recently suggested in one family.[39] Interestingly, similar variants in the kinesin motor domain of the *KIF1A* paralog *KIF5A* are involved in pure autosomal dominant SPG10. The nonsense variant in patient 23, which is not in the kinesin domain, is more difficult to understand.

Three other variants identified in the cohort of HSP patients (Supplementary File 5) were considered to be likely noncausative, as they did not segregate with the disease (variant in *KCNC3*), were not supported by biochemical analyses (*GALC*), or because no second pathogenic variant in the same gene was identified (*PNPLA6*).

DISCUSSION

Clinical exome sequencing in CA and HSP has a high diagnostic yield

Sequencing of single genes in monogenic disorders with high locus and allelic heterogeneity, such as CA and HSP, has a low diagnostic yield per gene. The most common dominant forms of CA are caused by polyglutamine-coding repeat expansions. Spinocerebellar ataxia 3 is the most prevalent CA in this group, accounting for 21% of all SCA patients worldwide, though large differences between populations have been reported.[6] The most common recessive type is Friedreich ataxia, with a prevalence of 2–4 in 100 000. As the total number of patients with recessive CA is unknown, it is as yet unclear which fraction is affected with Friedreich ataxia. The detection rate for other genotypes declines steeply, with many forms being extremely rare with prevalences below 1%. Once the ataxias with repeat expansion have been excluded, ~70 other genes associated with various other forms of ataxia remain candidates. Similarly, the most common form of HSP is autosomal dominant (AD) SPG type 4,[40] explaining ~30 to 40% of AD HSP. Many other HSP types are very rare, and sometimes have been described in only one or a few families. The number of loci involved in HSP has now exceeded 70. Thus, a generic test for parallel analyses of all the nonrepeat CA and HSP genes is expected to have a better overall yield than single gene tests. Considering the phenotypic overlap between CA and HSP, a combined analysis of genes involved in both groups of disorders is justified.[41] Indeed, clinical exome sequencing in this cohort showed a diagnostic yield (~1 in 4), comparable to earlier exome sequencing studies for CA.[4] The reasons for not detecting causative variants in approximately three-fourths of the patients in the cohort may be heterogeneous, including that the variants reside outside the tested regions (such as mitochondrial DNA, introns, and promoters), that the variants are in 'novel' genes yet to be identified as involved in a central motor disorder, that the variants are not being picked up by this test (exonic regions that are not targeted, poor coverage, or repeat expansions), or a nongenetic basis of the disorder. In addition, it is worthwhile noting that this was a selection of patients negative for conventional gene testing (4.5 genes on average), decreasing the *a priori* chance to find a variant by exome sequencing. Of note, a limitation of our study was the lack of systematicity in which genes were excluded before exome sequencing, preventing making reliable frequency estimates.

However, our selection of cases for this testing strategy, that is, a genetic etiology had to be suspected, might have led to a higher yield than when a less selected cohort of CA and HSP patients would have been tested. Future re-evaluation of the data (to detect variants in newly identified genes) or future retesting (ie, genome sequencing or mitochondrial DNA sequencing) may eventually provide a molecular cause for some of the patients in the 'unsolved' group.

Of note, before exome sequencing was undertaken, repeat expansion disorders were excluded in most CA patients (both here and the study from Nemeth *et al*[1]), and similarly *SPAST* and *SPG7* variants were excluded in most of the HSP patients. As it is still not possible to reliably detect repeat expansions by exome sequencing in CA, it remains necessary to test for these first. For HSP, however, clinical exome sequencing could be the first tier testing strategy that would increase its diagnostic yield.

Uncertainty of the pathogenicity of a subset of variants

One of the challenges in clinical genetics is the determination of the pathogenicity of a given variant. In part, variants are clearly or highly likely to be pathogenic based on functional studies, segregation analyses, frequencies of variants in patient and control sets, or simply on the nature of the variant (such as nonsense variant in haploinsufficiency diseases). Others may seem pathogenic, but proof is lacking, and their role in disease should be considered with great caution, and this has been stressed for variants with uncertain pathogenicity in ataxia by a discussion in the recent literature.[42,43] Therefore, such variants with uncertain pathogenicity (as in Tables 1b and 2b) should be considered carefully and not be used in clinical decision making. However, sharing these variants either in databases or via scientific papers is key to build the publically available data allowing scrutinization of the putative role of these variants in disease.

Clinical exome sequencing reveals 'old' gene–'new' disease associations

Clinical exome sequencing and filtering for a broader panel of genes involved in motor disorders has revealed some novel gene-to-disease associations. SPG7 is the most common form of autosomal recessive HSP and is in half of the cases associated with ataxia. Here, two patients (1 and 2) with a pure ataxia were found to be compound heterozygous for two common *SPG7* gene variants. Similarly, other exome sequencing studies had also revealed one patient with pure CA due to *SPG7* variants.[4,44] Thus, the clinical spectrum of *SPG7* variants is expanded to pure CA. Similarly, a homozygous variant in the *TH* gene in two siblings with HSP strongly suggests a role for this gene not only in dopamine-responsive dystonia, but also in HSP. A third observation is that of a dominant inheritance pattern of pure HSP in two independent families with *KIF1A* variants that were until now implicated in autosomal recessive HSP in two families or as *de novo* variants in severe intellectual disability and a motor phenotype.[37,38] For *SPG7*, this new association seems to be established with five families (this and other reports). For *TH* and *KIF1A*, the identification of other families and/or functional studies are warranted to establish their role in a wider disease spectrum.

Other observations in our cohort substantiate previously suggested 'allelic' disorders. These include the role of *PMM2* not only in a severe congenital disorder of glycosylation, but also in late-onset ataxia (patient 8), a second *VCP* variant (patient 18) that likely explains HSP and that suggest this form of HSP is allelic to ALS14 and IBMPFD1, and a possible second case of autosomal recessive spastic-ataxia-neuropathy syndrome (SPAX5)[45] caused by variants in *AFG3L2* (patient 24).

Some of our other findings cautiously hint toward other gene–disease associations, but clearly require confirmation by other cases or functional tests. These are the possible role of *LRRK2* (involved in Parkinson's disease) variants in HSP (patient 20), or that of *OPA1* (the gene for dominant optic atrophy plus syndrome) in neurological disorders without optic atrophy (patients 10 and 25).

Clinical exome sequencing uncovers unanticipated disorders

Other genetic disorders may escape from being recognized clinically, simply because they are extremely rare, or because the disease may not have progressed to the recognizable full-blown phenotype. For example, patient 5 with myoclonic epilepsy type 6 due to *GOSR2* variants had indeed gradually developed the classical phenotype of ataxia with myoclonic epilepsy and other seizure types, although at age 13 there is still no scoliosis. Because of the locus heterogeneity in ataxia–epilepsy syndromes and the rare occurrence of this rather new type, myoclonic epilepsy type 6 was not considered clinically. Another example in our cohort was that of Kufor–Rakeb syndrome (patient 17), who was initially diagnosed with a complex form of HSP because spastic tetraplegia was predominant. Only after establishing a molecular diagnosis, clinical re-examination ('reverse phenotyping') confirmed Kufor–Rakeb syndrome, leading to a favorable levodopa trial. Moreover, this case illustrates that unexpected exome sequencing findings can affect therapeutic management.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We acknowledge the Genome Technology Center at the Radboudumc and the Beijing Genome Institute Europe (Copenhagen, Denmark) for providing exome sequencing services. Dr van de Warrenburg receives research funds from the Radboud University Medical Center, BBMRI-NL (CP2013-57), the Gossweiler Foundation, and ZonMW. Dr Gilissen has received a 'VENI' grant from the Netherlands Organization for Scientific Research (Grant Number 916-14-043). Dr Scheffer receives research funds from the EU Framework Programme 7 (TECHGENE project; Health-F7-2009-223143).

1 Nemeth AH, Kwasniewska AC, Lise S *et al*: Next generation sequencing for molecular diagnosis of neurological disorders using ataxias as a model. *Brain* 2013; **136**: 3106–3118.
2 Ohba C, Osaka H, Iai M *et al*: Diagnostic utility of whole exome sequencing in patients showing cerebellar and/or vermis atrophy in childhood. *Neurogenetics* 2013; **14**: 225–232.
3 Novarino G, Fenstermaker AG, Zaki MS *et al*: Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. *Science* 2014; **343**: 506–511.
4 Fogel BL, Lee H, Deignan JL *et al*: Exome sequencing in the clinical diagnosis of sporadic or familial cerebellar ataxia. *JAMA Neurol* 2014; **71**: 1237–1246.
5 Neveling K, Feenstra I, Gilissen C *et al*: A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. *Hum Mutat* 2013; **34**: 1721–1726.
6 Bird TD: Hereditary ataxia overview; in: Pagon RA, Adam MP, Ardinger HH *et al* (eds): GeneReviews(R): Seattle, WA, 1993.
7 Fink JK: Hereditary spastic paraplegia overview; in: Pagon RA, Adam MP, Ardinger HH *et al* (eds): GeneReviews(R): Seattle, WA, 1993.
8 Graves TD, Imbrici P, Kors EE *et al*: Premature stop codons in a facilitating EF-hand splice variant of CaV2.1 cause episodic ataxia type 2. *Neurobiol Dis* 2008; **32**: 10–15.
9 Corbett MA, Schwake M, Bahlo M *et al*: A mutation in the Golgi Qb-SNARE gene GOSR2 causes progressive myoclonus epilepsy with early ataxia. *Am J Hum Genet* 2011; **88**: 657–663.
10 Figueroa KP, Minassian NA, Stevanin G *et al*: KCNC3: phenotype, mutations, channel biophysics-a study of 260 familial ataxia patients. *Hum Mutat* 2010; **31**: 191–196.
11 Goizet C, Boukhris A, Mundwiller E *et al*: Complicated forms of autosomal dominant hereditary spastic paraplegia are frequent in SPG10. *Hum Mutat* 2009; **30**: E376–E385.
12 Johnson J, Paisan-Ruiz C, Lopez G *et al*: Comprehensive screening of a North American Parkinson's disease cohort for LRRK2 mutation. *Neurodegener Dis* 2007; **4**: 386–391.
13 Delettre C, Griffoin JM, Kaplan J *et al*: Mutation spectrum and splicing variants in the OPA1 gene. *Hum Genet* 2001; **109**: 584–591.
14 Matthijs G, Schollen E, Pardon E *et al*: Mutations in PMM2, a phosphomannomutase gene on chromosome 16p13, in carbohydrate-deficient glycoprotein type I syndrome (Jaeken syndrome). *Nat Genet* 1997; **16**: 88–92.
15 Matthijs G, Schollen E, Heykants L, Grunewald S: Phosphomannomutase deficiency: the molecular basis of the classical Jaeken syndrome (CDGS type Ia). *Mol Genet Metab* 1999; **68**: 220–226.
16 Blanchard MG, Willemsen MH, Walker JB *et al*: De novo gain-of-function and loss-of-function mutations of SCN8A in patients with intellectual disabilities and epilepsy. *J Med Genet* 2015; **52**: 330–337.
17 McDermott CJ, Dayaratne RK, Tomkins J *et al*: Paraplegin gene analysis in hereditary spastic paraparesis (HSP) pedigrees in northeast England. *Neurology* 2001; **56**: 467–471.
18 Brugman F, Scheffer H, Wokke JH *et al*: Paraplegin mutations in sporadic adult-onset upper motor neuron syndromes. *Neurology* 2008; **71**: 1500–1505.
19 Ng PC, Henikoff S: Predicting deleterious amino acid substitutions. *Genome Res* 2001; **11**: 863–874.
20 Casari G, Marconi R: Spastic paraplegia 7; in: Pagon RA, Adam MP, Ardinger HH *et al* (eds): GeneReviews(R): Seattle, WA, 1993.
21 van Gassen KL, van der Heijden CD, de Bot ST *et al*: Genotype-phenotype correlations in spastic paraplegia type 7: a study in a large Dutch cohort. *Brain* 2012; **135**: 2994–3004.
22 O'Brien JE, Meisler MH: Sodium channel SCN8A (Nav1.6): properties and de novo mutations in epileptic encephalopathy and intellectual disability. *Front Genet* 2013; **4**: 213.
23 Boisse Lomax L, Bayly MA, Hjalgrim H *et al*: 'North Sea' progressive myoclonus epilepsy: phenotype of subjects with GOSR2 mutation. *Brain* 2013; **136**: 1146–1154.
24 Vega AI, Perez-Cerda C, Abia D *et al*: Expression analysis revealing destabilizing mutations in phosphomannomutase 2 deficiency (PMM2-CDG): expression analysis of PMM2-CDG mutations. *J Inherit Metab Dis* 2011; **34**: 929–939.
25 Vermeer S, Kremer HP, Leijten QH *et al*: Cerebellar ataxia and congenital disorder of glycosylation Ia (CDG-Ia) with normal routine CDG screening. *J Neurol* 2007; **254**: 1356–1358.
26 Coman D, McGill J, MacDonald R *et al*: Congenital disorder of glycosylation type 1a: three siblings with a mild neurological phenotype. *J Clin Neurosci* 2007; **14**: 668–672.

27 Bonifert T, Karle KN, Tonagel F *et al*: Pure and syndromic optic atrophy explained by deep intronic OPA1 mutations and an intralocus modifier. *Brain* 2014; **137**: 2164–2177.
28 Slabicki M, Theis M, Krastev DB *et al*: A genome-scale DNA repair RNAi screen identifies SPG48 as a novel gene associated with hereditary spastic paraplegia. *PLoS Biol* 2010; **8**: e1000408.
29 Schlipf NA, Schüle R, Klimpe S *et al*: AP5Z1/SPG48 frequency in autosomal recessive and sporadic spastic paraplegia. *Mol Genet Genom Med* 2014; **2**: 4.
30 Fuger P, Sreekumar V, Schule R *et al*: Spastic paraplegia mutation N256S in the neuronal microtubule motor KIF5A disrupts axonal transport in a Drosophila HSP model. *PLoS Genet* 2012; **8**: e1003066.
31 Montenegro G, Rebelo AP, Connell J *et al*: Mutations in the ER-shaping protein reticulon 2 cause the axon-degenerative disorder hereditary spastic paraplegia type 12. *J Clin Invest* 2012; **122**: 538–544.
32 de Bot ST, Schelhaas HJ, Kamsteeg EJ, van de Warrenburg BP: Hereditary spastic paraplegia caused by a mutation in the VCP gene. *Brain* 2012; **135**: e223, author reply e224.
33 Guyant-Marechal L, Laquerriere A, Duyckaerts C *et al*: Valosin-containing protein gene mutations: clinical and neuropathologic features. *Neurology* 2006; **67**: 644–651.
34 Willemsen MA, Verbeek MM, Kamsteeg EJ *et al*: Tyrosine hydroxylase deficiency: a treatable disorder of brain catecholamine biosynthesis. *Brain* 2010; **133**: 1810–1822.
35 Fan Z, Greenwood R, Felix AC *et al*: GCH1 heterozygous mutation identified by whole-exome sequencing as a treatable condition in a patient presenting with progressive spastic paraplegia. *J Neurol* 2014; **261**: 622–624.
36 Klebe S, Lossos A, Azzedine H *et al*: KIF1A missense mutations in SPG30, an autosomal recessive spastic paraplegia: distinct phenotypes according to the nature of the mutations. *Eur J Hum Genet* 2012; **20**: 645–649.
37 Hamdan FF, Gauthier J, Araki Y *et al*: Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am J Hum Genet* 2011; **88**: 306–316.
38 Lee JR, Srour M, Kim D *et al*: De novo mutations in the motor domain of KIF1A cause cognitive impairment, spastic paraparesis, axonal neuropathy, and cerebellar atrophy. *Hum Mutat* 2015; **36**: 69–78.
39 Ylikallio E, Kim D, Isohanni P *et al*: Dominant transmission of de novo KIF1A motor domain variant underlying pure spastic paraplegia. *Eur J Hum Genet* 2015; **23**: 1427–1430.
40 Durr A, Tallaksen C, Depienne C: Spastic paraplegia 4; in: Pagon RA, Adam MP, Ardinger HH *et al* (eds): GeneReviews(R): Seattle, WA, 1993.
41 de Bot ST, Willemsen MA, Vermeer S, Kremer HP, van de Warrenburg BP: Reviewing the genetic causes of spastic-ataxias. *Neurology* 2012; **79**: 1507–1514.
42 Sandford E, Li JZ, Burmeister M: Evaluation of exome sequencing variation in undiagnosed ataxias. *Brain* 2015; **138**: e383.
43 Pyle A, Griffin H, Keogh MJ, Horvath R, Chinnery PF: Reply: Evaluation of exome sequencing variation in undiagnosed ataxias. *Brain* 2015; **138**: e384.
44 Pyle A, Smertenko T, Bargiela D *et al*: Exome sequencing in undiagnosed inherited and sporadic ataxias. *Brain* 2015; **138**: 276–283.
45 Pierson TM, Adams D, Bonn F *et al*: Whole-exome sequencing identifies homozygous AFG3L2 mutations in a spastic ataxia-neuropathy syndrome linked to mitochondrial m-AAA proteases. *PLoS Genet* 2011; **7**: e1002325.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)